ALASKA OREGON WASHINGTON

PACIFIC NORTH WEST FOREST AND RANGE EXPERIMENT STATION

USDA FOREST SERVICE RESEARCH NOTE

PNW-162 *August 1971*

3 P SAMPLE LOG SCALING

by

Floyd A. Johnson, *Biometrician*
Pacific Northwest Forest & Range Experiment Station

James B. Lowrie, *Measurement Specialist*
Pacific Northwest Region

and

Martin Gohlke, *Check Scaler*
Pacific Northwest Region

ABSTRACT

An application of the 3P sample selection procedure to sample log scaling is described in this report along with results from three tests. The method looks promising for some situations, and it may prove superior to the sample log scaling methods now in use.

Keywords: Log scaling, sample designs (forestry), timber estimating, defect deduction (merchantable volume), forest measurement.

INTRODUCTION

The 3P sample selection procedure,[1] which was originally developed for timber cruising, has been applied to log scaling with results which appear to be promising. This application was tested first in Idaho[2] and later in Oregon. Results of the Oregon test will be described in this report along with details of procedure.

Under this procedure the scaler quickly guesses the small-end diameter and length of each log in the group of logs for which an estimate of total volume is required. The two guesses are then translated to gross log volume in terms of the Scribner Decimal C log rule, and "guessed" gross volume is compared with a random number from a specially prepared set of random numbers.

If the random number is greater than guessed gross volume or if the random number is zero, the guess is merely recorded and the log is passed by. If the random number is less than or equal to the guess, but not zero, the log is carefully scaled for both gross and net volume.

The largest number in the set of random numbers should be at least as large as the largest guessed log volume. Since each number from one to the largest will tend to be equally frequent in the set, the number of random numbers which will qualify any log for the sample will tend to be proportional to the guessed volume of that log. This follows from the sample selection rule given in the previous paragraph. If the set were infinitely large, the number frequencies would be exactly equal and the proportionality would be absolute. In other words, there will be about twice as many numbers available for qualifying a log with a predicted volume of, say, 400 board feet as for a log with a predicted volume of 200 board feet. A 1,000-board-foot log will tend to have 10 times as many qualifying numbers, and therefore 10 times as many chances of getting into the sample as will a 100-board-foot log. The frequency of zeros in the set of random numbers will be more or less large depending on the desired sample size.

Obviously, large logs are favored under 3P sample selection, and this must be compensated for. The compensation is accomplished by taking the

---

1/ L. R. Grosenbaugh. Some suggestions for better sample tree measurement. Proc. Soc. Amer. Forest. Meet. Oct. 20-23, 1963, Boston, Mass., p. 36-42, 1964.

2/ Speech presented by James C. Space, Supervisory Forester, Sandpoint Ranger District, Kaniksu National Forest, West. Pine Reg. Log Scaling Soc. Ann. Meet., Coeur d'Alene, Idaho, Oct. 9, 1969.

ratio of scaled volume to guessed volume for a sample log as the observation which has unit weight, and then multiplying the average of all such ratios by the sum of the guessed volumes for all logs in the group.

Thus

$$V = \left(\sum^{n} R/n\right)\left(\sum^{N} KPI\right)$$

where V = estimated total scaled volume (net or gross) for a group of logs,

KPI = guessed volume of one log by the Scribner Decimal C log rule.

$\sum^{N} KPI$ = sum of guessed volumes over all logs in the group,

N = number of logs in the group,

R = measured volume divided by guessed volume for one sample log,

$\sum^{n} R$ = sum of R over all sample logs in the group,

and n = number of sample logs in the group.

If the groups of logs being sampled by one set of random numbers have been divided into subgroups (strata), the above equation will apply only to the logs in a single stratum. In that event, estimated total volume for the entire group of logs (all strata) is obtained simply by adding estimated stratum volumes.

THREE TESTS

Three tests of 3P sample log scaling were made recently. In two of these tests, the logs were in rafts and in the third the logs were on trucks. Scaling procedures used on all three tests were those prescribed for National Forest timber sales in western Oregon and western Washington. There were three species in the truck test, and each species was identified as a separate subgroup or stratum. Each of the two raft tests consisted of a single stratum.

RANDOM NUMBERS

Random numbers created for the truck test are illustrated in figure 1. These numbers are comparable to guessed log volumes by the Scribner Decimal C log rule. For example, a log with a 300-board-foot guessed volume would

Figure 1

RANDOM NUMBERS FOR TRUCK TEST PAGE 1 07/29/71
L= 762522741, LIM= 1500, K= 300, KZ= 440 TRUK

0	223	62	131	1	270	239	0	209	77
271	0	0	0	0	0	0	17	10	0
232	177	0	141	128	13	141	36	15	0
0	0	61	78	213	182	140	41	98	283
0	293	294	78	128	136	208	141	0	78
0	10	0	295	41	83	209	0	0	0
83	39	0	0	0	96	170	0	0	45
0	134	222	98	10	150	102	57	0	3
110	204	0	67	41	74	119	0	19	114
186	183	241	201	0	0	162	213	0	124
270	126	32	0	118	0	0	232	0	0
0	54	0	24	0	198	0	63	253	0
292	0	54	110	0	0	25	162	116	0
219	132	0	237	202	0	223	129	0	0
72	224	113	18	46	0	124	118	0	0
0	0	0	177	0	91	252	107	140	163
167	0	196	0	0	0	198	282	0	116
0	55	224	0	256	7	156	0	255	0
0	0	158	68	0	149	286	90	70	0
103	252	0	203	186	145	83	0	87	267
107	257	187	80	256	217	82	0	70	0
0	101	240	201	0	223	67	0	174	123
90	31	0	218	0	46	0	0	129	150
183	0	142	0	0	87	77	266	0	0
0	233	42	0	233	120	0	0	77	194
202	144	0	245	0	27	0	107	149	150
0	0	0	139	20	0	166	239	10	0
0	137	61	68	145	143	0	0	0	0
0	221	5	206	146	0	6	26	246	179
0	0	102	243	193	194	143	0	28	156
0	0	0	231	118	161	0	190	93	0
164	207	0	298	0	277	192	0	198	135
278	44	0	0	257	197	0	274	0	0
214	293	288	6	0	157	0	0	10	51
0	0	120	0	0	124	292	283	239	145
43	0	244	45	100	298	148	0	286	139
0	292	222	150	0	0	67	0	238	0
95	75	111	254	0	197	201	162	88	298
252	217	0	114	182	274	0	179	33	247
0	96	257	285	116	0	148	23	152	191
0	0	0	136	71	93	260	0	45	0
29	25	51	96	15	0	0	0	12	0
0	138	0	0	130	298	125	48	49	0
62	158	44	288	0	0	0	106	177	10
197	32	91	0	0	0	50	10	240	156
137	0	0	0	266	84	67	0	140	114
0	173	239	123	107	0	224	109	2	0
1	0	56	0	136	160	298	46	7	0
165	203	63	234	118	192	218	170	31	284
0	219	146	131	0	0	167	111	71	0

46983
TRUK TRUK TRUK TRUK TRUK TRUK TRUK TRUK TRUK TRUK 327

be coded 30, and any random number from 1 through 30 would qualify that log for the sample of precisely measured logs. The coded guessed volume of a log will be called *KPI* following Grosenbaugh's notation.

A computer program called THRP[3/] is used to create the random number sets. This program requires only three items of information:

1. An advance estimate (K) of the gross volume of the largest log in the entire group of logs that is to be sampled (i.e., including all strata) with one set of numbers. This advance estimate cannot and need not be precise. An overestimate is better than an underestimate. If the timber to be scaled had been previously cruised, the cruise data are a likely source for this advance estimate.

2. An advance estimate of the total gross volume in all of the strata which are to be sampled with one set of random numbers. Information from a timber sale cruise could be used for this advance estimate. This advance estimate is an attempt to anticipate the eventual sum of the guessed volumes $(\sum^{N} KPI)$ over all logs in all strata. If an estimate is inaccurate, the only penalty will be that the sampling error for the eventual estimate of total volume (all strata) will be either greater or less than desired.

3. An advance estimate of the total number of logs in all strata. This merely tells the computer how many numbers to produce. Since the numbers are inexpensive, there is no harm in using a rather large overestimate.

Item 2 must be divided by an arbitrary value for n (i.e., expected number of sample logs in all strata that will be sampled with one set of random numbers). The quotient (called KZ) is then punched on a THRP control card along with items 1 and 3. The arbitrary value for n will depend on the level of precision (i.e., size of the standard error) required for the eventual estimates of total sale volume (all strata) and on the expected average within-stratum coefficient of variation for R, the ratio of measured to guessed volume. Initially, the choice of n will be difficult, but it will become easier as experience accumulates over many samplings.

[3/] L. R. Grosenbaugh. Three-pee sampling theory. USDA Forest Serv. Res. Paper PSW-21. Berkeley, Calif., Pac. Southwest Forest & Range Exp. Sta., 53 p., 1965.

DATA

The tally form and tallied information for 31 of the 266 logs in the truck test are shown in figure 2 to help describe the 3P procedure.

Thirteen spaces in the heading of the tally form have been allocated to an identification of the entire group of logs which are being sampled with a single set of random numbers. A single job could be an entire timber sale, or any portion of a timber sale (e.g., logs sold in a month or a quarter), or a single raft. At least one of these spaces must be used with a distinctive code.

Information from each log in a timber sale will be recorded on a single row in the body of the form. The word "log" as used here refers to any whole log 40 feet or under and to scaled segments of whole logs which are longer than 40 feet.

Columns headed "log identification" may be left blank, or they could be used for a truck number, log number, or any remarks pertinent to an individual log that a scaler might care to make. Entries in these columns will not be used in the subsequent processing of data.

A distinctive stratum code must be used to identify the logs for which separate estimates of total volume are required. Usually each species will constitute a stratum; but a stratum can include several species, and there can be more than one stratum per species (e.g., sound and defective Douglas-fir). Stratum codes can be any combination of alpha or numeric characters.

A code = in the third column means that the log in question will not be part of the sample and will be scaled without question. It will be called a "sure" log following Grosenbaugh. This would be appropriate whenever there are only a few logs in a stratum or whenever the logs in a stratum are of such high value that any amount of sampling error would be intolerable. An asterisk (*) in this column means that the log in question is part of a group of logs being sampled and has been selected for precise measurement. If the third column is blank, the log is part of the group of logs being sampled but was not selected.

Columns headed *KPI* are for the coded guessed volumes of those logs which are sampled. There must be an entry in these columns whenever there is an asterisk (*) in the preceding column or whenever the preceding column is blank. These columns can be left blank whenever the code = appears in the preceding column. However no harm will be done if a *KPI* is entered for a "sure" log because it will be ignored by the computer. If the *KPI* for any log happens to be larger than *K*, the advance estimate of largest *KPI* that was used

Figure 2

3P SAMPLE LOG SCALING
SALE IDENTIFICATION

TRUCK TEST

LOG IDENTIFICATION	STRATUM	* =	KPI	GROSS L	GROSS D	NET L	NET D
4656	DF		86				
4656	DF		53				
4656	DF	*	186	40	34	32	33
4656	DF		15				
4656	DF		65				
4656	DF		129				
4656	DF		13				
4656	DF		56				
4656	DF		49				
2118	DF		6				
2118	NF		7				
2118	NF		5				
2118	DF	*	76	40	18	40	18
2118	DF		10				
2118	DF		9				
2118	DF		15				
2118	DF		7				
2118	DF		29				
2118	DF		12				
2118	DF		7				
2118	DF		70				
2118	DF		5				
2118	DF		20				
2118	DF		34				
2118	DF		32				
2118	DF		49				
4304	DF		53				
4304	DF	*	53	36	18	36	18
4304	DF		40				
4304	DF		66				
4304	DF		23				

to create the random numbers, the scaler may either reduce this *KPI* to *K* before comparing it with a random number, or he may call the log, "sure."

Gross and net lengths (L) and diameters (D) are for individual logs or for segments of whole logs which are over 40 feet long. Each segment of a whole log which is over 40 feet will be treated separately in the sample selection process. Thus entries will be made on two rows of the tally form for a two-segment-long log, and either one or both of these segments may become part of the sample. A sample log will be distinguished from a nonsample log by entries in the L and D columns. All "sure" logs will, of course, have entries in the L and D columns.

Defect deductions leading to the lengths and diameters in the net columns of the tally form were made according to National Forest scaling practice for western Oregon and western Washington.

Twenty-eight of the 31 logs listed in figure 2 are obviously nonsample logs, since nothing is recorded for them in the L and D columns. There were no "sure" logs in the truck test.

Information on the tally forms is punched on cards with one card per row of information. Cards are then sorted by stratum. Decks for several strata and for several jobs can be stacked and processed in one computer run by a program called SLOT which was written for the CDC 6400 computer.

OUTPUT SPECIFICATIONS

Output from SLOT is in two parts. The first part (see fig. 3) is simply a listing of information on individual sample logs and on individual "sure" logs, if any. There will be a separate listing for each stratum. Figure 3 shows only the listing for the Douglas-fir stratum of the truck test.

A portion of the information listed is simply a relisting of input data from the tally form, but gross and net Scribner volumes, as selected by the computer, are also shown along with the ratios (R) of measured to guessed volumes and with an expansion factor for each log. If the volume of a log is multiplied by this factor and if the products of this multiplication for all sample logs in a stratum are added, the result will be estimated total volume for the sampled portion of that stratum. A simple sum of the volumes of "sure" logs in a stratum added to the estimated total volume for the sampled portion of the stratum will equal the estimated total volume in the entire stratum.

Figure 3

S L O T 3P LOG SCALING WEST SIDE

INFORMATION ON 3P AND SURE LOGS BY STRATUM 07/21/71

STRATUM IDENT DF PAGE 1 NUMBER OF KPI LOGS 224

SALE IDENT TRUCK TEST

LOG IDENTIFICATION	KPI	GROSS L	GROSS D	NET L	NET D	VOLUME GROSS	VOLUME NET	DEFECT VOL.	R GROSS	R NET	EXPANSION FACTOR
4656	186	40	34	32	33	200	157	43	1.075	.844	2.573
2118	76	40	18	40	18	53	53	0	.697	.697	6.298
4304	53	36	18	36	18	48	48	0	.906	.906	9.031
3610	24	16	20	16	17	28	18	10	1.167	.750	19.944
4162	84	40	20	36	20	70	63	7	.833	.750	5.698
3942	60	40	19	40	16	60	40	20	1.000	.667	7.978
3942	19	40	14	40	13	29	24	5	1.526	1.263	25.193
2796	106	34	24	34	23	86	80	6	.811	.755	4.516
2796	100	20	34	20	31	100	89	11	1.000	.890	4.787
2796	61	17	31	8	31	75	36	39	1.230	.590	7.847
6639	49	40	17	40	17	46	46	0	.939	.939	9.769
4656	18	34	13	34	12	21	17	4	1.167	.944	26.593
1024	89	24	21	24	18	46	32	14	.517	.360	5.378
1024	200	40	34	40	32	200	184	16	1.000	.920	2.393
1025	142	40	27	20	27	137	68	69	.965	.479	3.371
1025	89	38	20	38	16	66	38	28	.742	.427	5.378
2798	76	38	22	34	20	79	60	19	1.039	.789	6.298
2798	51	34	19	34	15	51	30	21	1.000	.588	9.386
1027	23	26	17	24	16	30	24	6	1.304	1.043	20.812
1027	103	36	24	36	20	91	63	28	.883	.612	4.647
1027	64	22	26	20	26	69	62	7	1.078	.969	7.479
1027	217	38	34	36	33	190	176	14	.876	.811	2.206
2121	176	36	33	34	33	176	167	9	1.000	.949	2.720
1028	300	40	40	38	38	300	254	46	1.000	.847	1.596

The equation for estimated total stratum volume (see page 3) can be rewritten to clarify the expansion factors for individual logs:

$$V = \left[y_1 \frac{\sum^{N} KPI}{KPI_1 n} + y_2 \frac{\sum^{N} KPI}{KPI_2 n} - - - - - - - - - - - - - + y_n \frac{\sum^{N} KPI}{KPI_n n} \right]$$

where y_i = scaled volume for an individual sample log.

Thus, the expansion factor for the first log in figure 3 can be calculated by taking ΣKPI for the DF stratum in figure 4 (11,488) and dividing it by the DF sample size in figure 4 (24) and again by the KPI for the first log (186):

$$\frac{11,488}{(24)(186)} = 2.573.$$

Part 2 of the SLOT output contains summary information for the entire sale (see fig. 4). The sums of KPI in the farthest right column of this output are for all logs (sample + nonsample). Equations used for calculating coefficients of variation and standard errors shown in figure 4 are given in the appendix. Estimated total volumes are expressed in units of 10 so the net estimate for Douglas-fir is 89,930 board feet. Defect percentages are simply differences between gross and net total volumes expressed as a percent of gross total volume.

Summary information for the two raft tests is shown in figures 5 and 6. There were three "sure" logs in one raft test (fig. 5). They were called "sures" because they appeared to be excessively defective. Had they been selected for the sample, they might have had very low volume/KPI ratios. This would have increased coefficient of variation and sampling error for the net volume estimates considerably. If there are likely to be many logs that are highly defective, it might be better to put them in a different stratum and to sample that stratum rather than scale all the logs as "sure." Thus, there could be a Douglas-fir stratum (DF) and a defective Douglas-fir stratum (DDF).

Lower coefficients of variation and therefore lower standard errors might have resulted if the KPI's had been in terms of net rather than gross log volume. However this could lead to bias since the scaler's judgments regarding defect in his scaling of the logs selected for the sample might be conditioned by the quick judgments he made while applying his KPI's. KPI and scaled volumes on individual logs must be independent.

Figure 4

S L O T 3P LOG SCALING WEST SIDE

SUMMARY FOR SALE 07/21/71

SALE IDENT TRUCK TEST PAGE 1

STRATUM		SAMPLE LOGS	GROSS VOL.DC	GROSS SE	GROSS CV	NET VOL.DC	NET SE	NET CV	DEFECT	TOTAL LOGS	SUM OF KPI
DF	3P	24	11371	4.27	20.93	8993	5.41	26.52	20.907	224	11488
	SURE	0	0			0			0.000	0	
	TOTAL		11371	4.27		8993	5.41		20.907	224	
WP	3P	3	291	13.21	22.88	287	14.65	25.37	1.258	10	329
	SURE	0	0			0			0.000	0	
	TOTAL		291	13.21		287	14.65		1.258	10	
NF	3P	3	982	12.36	21.41	787	11.51	19.93	19.924	32	1224
	SURE	0	0			0			0.000	0	
	TOTAL		982	12.36		787	11.51		19.924	32	
SALE TOTAL			12644	3.97	21.28	10067	4.94	26.21			

Figure 5

S L O T 3P LOG SCALING WEST SIDE

SUMMARY FOR SALE 07/26/71

SALE IDENT RAFT 639C PAGE 1

STRATUM		SAMPLE LOGS	GROSS VOL.DC	GROSS SE	GROSS CV	NET VOL.DC	NET SE	NET CV	DEFECT	TOTAL LOGS	SUM OF KPI
DF	3P	31	14188	3.00	16.69	13286	3.61	20.09	6.356	238	14442
	SURE	3	184			39			78.804	3	
	TOTAL		14372	2.96		13325	3.60		7.284	241	
SALE TOTAL			14372	2.96	16.69	13325	3.60	20.09			

Figure 6

S L O T 3P LOG SCALING WEST SIDE

SUMMARY FOR SALE 07/26/71

SALE IDENT RAFT 61 PAGE 1

STRATUM		SAMPLE LOGS	GROSS VOL.DC	GROSS SE	GROSS CV	NET VOL.DC	NET SE	NET CV	DEFECT	TOTAL LOGS	SUM OF KPI
DF	3P	35	35168	1.77	10.49	30026	2.88	17.03	14.621	162	34807
	SURE	0	0			0			0.000	0	
	TOTAL		35168	1.77		30026	2.88		14.621	162	
SALE TOTAL			35168	1.77	10.49	30026	2.88	17.03			

Normally, a single set of random numbers will be used for all strata in a sale, but separate sets could be used for individual strata if more flexibility in sample sizes is essential. However, several sets of numbers are awkward to handle.

SUMMARY

The coefficients of variation (CV) in figures 4, 5, and 6 suggest that 3P log scaling may be very efficient in the sense of providing precise estimates of total volume at small cost. For example, CV's around 20 percent can be expected; with a 20-percent CV, 100 individual sample logs would give an estimate of total timber sale volume which has a 2-percent standard error. Under sample scaling, where the entire truckload of logs is the sampling unit rather than the individual log, CV's of 20 percent or greater are quite common; but a 20-percent CV in this case would require 100 truckloads of logs instead of 100 individual logs for a 2-percent standard error. The cost of scaling 100 truckloads of logs cannot, of course, be compared directly with the cost of scaling 100 individual logs, because the 3P method also requires a guessed volume for every log in the sale. However, in spite of this additional expense, it seems that 3P sample scaling may have a decided advantage.

APPENDIX

Stratum coefficients of variation are standard deviations of the individual sample of R values expressed as a percent of mean R. For "sale total," these CV's are weighted averages over all strata. Thus

$$CV_S = 100 \frac{S}{\bar{R}}$$

where $S = \sqrt{\frac{\Sigma R^2 - (\Sigma R)^2/n}{n-1}},$

$\bar{R} = \Sigma R/n,$

and CV_S = stratum coefficient of variation;

and

$$CV_T = \frac{100\Sigma n}{\Sigma(\bar{R}n)}\sqrt{\frac{\Sigma S^2(n-1)}{\Sigma n-k}}$$

where k = number of strata,

CV_T = average coefficient of variation for all strata.

Standard errors in percent ($SE\%$) for volume in the 3P portion of any stratum can be calculated as follows:

$$SE\%_S = CV/\sqrt{n}$$

If there is "sure" volume in a stratum, the standard error of stratum total volume results from multiplying the standard error of the 3P portion by the ratio of 3P volume to stratum total volume. Thus, for gross volume in figure 5:

$$SE\%_{(Total)} = 3.00(14{,}188/14{,}372)$$

$$= 2.96$$

The standard error in percent for total volume for all strata is calculated by this equation:

$$SE\%_T = \frac{\sqrt{\Sigma(SE\%V)^2}}{\Sigma V}$$

where V = stratum total volume (summations are for all strata).

The mission of the PACIFIC NORTHWEST FOREST AND RANGE EXPERIMENT STATION is to provide the knowledge, technology, and alternatives for present and future protection, management, and use of forest, range, and related environments.

Within this overall mission, the Station conducts and stimulates research to facilitate and to accelerate progress toward the following goals:

1. Providing safe and efficient technology for inventory, protection, and use of resources.
2. Development and evaluation of alternative methods and levels of resource management.
3. Achievement of optimum sustained resource productivity consistent with maintaining a high quality forest environment.

The area of research encompasses Oregon, Washington, Alaska, and, in some cases, California, Hawaii, the Western States, and the Nation. Results of the research will be made available promptly. Project headquarters are at:

College, Alaska	Portland, Oregon
Juneau, Alaska	Roseburg, Oregon
Bend, Oregon	Olympia, Washington
Corvallis, Oregon	Seattle, Washington
La Grande, Oregon	Wenatchee, Washington